확 률 론

#1 어떤 사람이 자기 집에서 회사까지 직접 운전하여 출근하는데 걸리는 시간(단위 : 분)을 확률변수 X 라고 할 때, X 는 [30,50] 에서 균등분포를 따른다고 한다. 그의 출근 시간은 8시까지이다. 그의 집에서 매일 아침 7시 15분에 출발한다고 할 때 지각할 확률을 구하시오.

$$P(X \geq 45) = 1 - P(X < 45) = 1 - \frac{15}{20} = \frac{1}{4}$$

#2 어떤 대학의 입시에 응시한 1,000명의 수학 성적을 조사하였더니 평균이 50점이고, 표준편차가 15인 정규분포를 이루었다. 이 때 점수가 35점부터 65점 사이인 학생 수를 구하라.

$X \sim N(50, 15^2)$,

$P(35 \leq X \leq 65) = P(-1 \leq Z \leq 1) = 0.682689$,

$0.683 \times 1{,}000 = 683$ 명

#3 어떤 대학의 강좌를 등록한 1,000명의 시험 성적은 평균이 65점이고, 표준편차가 10인 정규분포를 따른다. 80점 받은 학생의 순위는?

$X \sim N(65, 10^2)$,

$P(X \geq 80) = P\left(Z \geq \frac{80-65}{10}\right) = P(Z \geq 1.5) = 0.067$,

$0.067 \times 1{,}000 = 67$ 등

#4 어떤 기계를 수리하는데 걸리는 시간은 $\lambda = 3$ 인 지수 분포를 따른다고 할 때 다음을 구하시오.

(1) 수리 시간이 30분을 초과할 확률

(2) 수리 시간이 5시간을 초과했을 때, 적어도 30분이 더 소요될 확률

(1) 수리 시간이 30분을 초과할 확률

$$P(X > 0.5) = \int_{0.5}^{\infty} 3e^{-3x}dx = [-e^{-3x}] = 0 - e^{-1.5} = e^{-1.5} = 0.2231$$

(2) 수리 시간이 5시간을 초과했을 때, 적어도 30분이 더 소요될 확률

지수 분포의 무기억성의 원리에 의해 $P(X > 5\text{시간 } 30\text{분} | X > \text{시간}) = P(X > 30\text{분})$ 이다.

$$P(X > 0.5) = \int_{0.5}^{\infty} 3e^{-3x}dx = [-e^{-3x}] = 0 - e^{-1.5} = e^{-1.5} = 0.2231$$

#5 어느 방송국의 주말 연속극 시청률이 20%였다. 새 프로그램 편성 후 100명의 시청자를 뽑아 연속극의 시청 여부를 조사하였다. 다음을 구하라.

(1) 이 연속극에 대한 시청률이 전과 동일하다면 100명 중 15명 이하가 시청할 확률

(2) (1)에서 이 연속극을 25명 이상 시청할 확률

$X \sim B(100, \frac{1}{5})$ 를 정규분포로 근사하면 $X \sim N(20, 4^2)$

(1) 이 연속극에 대한 시청률이 전과 동일하다면 100명 중 15명 이하가 시청할 확률

$P(X \leq 15) = P\left(Z \leq \frac{15-20}{4}\right) = P(Z \leq -1.125) = 0.13$

(2) (1)에서 이 연속극을 25명 이상 시청할 확률

$P(X \geq 25) = P\left(Z \geq \frac{25-20}{4}\right) = P(Z \geq 1.125) = 0.13$

#6 어떤 강철 막대를 부러뜨리는데 필요한 최소의 힘의 평균은 12432이고, 표준편차가 25인 정규분포를 따른다고 한다. 400개의 샘플을 뽑았을 때 최소한 349개가 12400의 힘에 부러지지 않을 확률을 구하라.

$P(X>12400) = P\left(Z>\frac{12400-12432}{25}\right) = P(Z>-1.28) = 0.9$ 이므로 막대 하나가 12400의 힘에 부러지지 않을 확률은 90% 이다.

12400의 힘으로 부러지지 않을 막대의 개수를 T 이라고 할 때, $T\sim B(400,\frac{9}{10})$ 를 따른다. 이를 정규분포로 근사화시키면 $T\sim N(360,6^2)$ 이다.

$P(T\geq 349) = P\left(Z\geq \frac{348.5-360}{6}\right) = P(Z\geq -1.9167) = 0.97$ 이므로 97%이다.

#7 어느 기기에서 사용하는 전지의 수명은 평균 3개월, 분산 1인 정규분포를 따른다. 최소 400개월간 계속 기기를 사용할 수 있는 확률이 0.9772가 넘게 하려면 적어도 몇 개를 구입해야 하는가?

1개 수명 :$X_1 \sim N(3,1^2)$

2개 수명 :$X_2 \sim N(3\times 2,1^2\times 2)$

3개 수명 :$X_3 \sim N(3\times 3,1^2\times 3)$

…

n개 수명 :$X_n \sim N(3n,n)$

$p(X_n > 40) > 0.9772$

$p(X_n > 40) > P(Z>-2)$

$p(Z> \frac{40-3n}{\sqrt{n}}) > P(Z>-2)$

$\frac{40-3n}{\sqrt{n}} \leq -2$

$n \geq 16$

적어도 16개를 구입해야 한다.

#8 X, Y는 결합확률질량함수가 $(x,y) = (0,1),(0,2),(1,2),(1,3)$ 에 대하여 $f_{X,Y}(x,y) = \frac{2x+y}{12}$ 이고, 그 외는 0인 확률변수들이다. $E(X^7Y^2+2XY)$ 는?

Y \ X	0	1
1	$\frac{1}{12}$	$\frac{0}{12}$
2	$\frac{2}{12}$	$\frac{4}{12}$
3	$\frac{0}{12}$	$\frac{5}{12}$

$$E(X^7Y^2+2XY) = \sum_{Y=1}^{3}\sum_{X=0}^{1}(X^7Y^2+2XY)f(x,y)$$

$$(1^7\times 2^2+2\times 1\times 2)\times\frac{4}{12} + (1^7\times 3^2+2\times 1\times 3)\times\frac{5}{12} = \frac{64}{12}+\frac{270}{12} = \frac{334}{12}$$

#9 X, Y는 결합확률밀두함수가 $(x,y) = 10xy^2$, $0<x<y<1$ 이고 그 외는 0인 확률변수들이다. X에 대한 주변확률밀도함수를 구하라.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$
$$= \int_x^1 10xy^2dy = 10x\int_x^1 y^2dy = 10x\left[\frac{y^3}{3}\right] = \frac{10x}{3}(1-x^3)$$

#10 X, Y가 독립이면 임의의 함수 g, h 에 대하여 다음이 성립한다.

$$E[g(x)h(y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_{X,Y}(x,y)dxdy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy = \left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\right) = E[g(x)]E[h(y)]$$

#11 X, Y가 분포가 다음과 같을 때 독립인지 조사하라. 또 $E[X^2Y^3]$, $E[X^2]$, $E[Y^3]$ 을 구하라.

Y \ X	0	1
1	$\frac{1}{12}$	$\frac{0}{12}$
2	$\frac{2}{12}$	$\frac{4}{12}$
3	$\frac{0}{12}$	$\frac{5}{12}$

$P(X=1,Y=1) = 0,\quad P(X=1) = \frac{9}{12},\quad P(Y=1) = \frac{1}{12},\quad \therefore P(X=1,Y=1) \neq P(X=1)P(Y=1)$

이므로 독립이 아니다.

$$E[X^2] = 0^2\frac{3}{12}+1^2\frac{9}{12} = \frac{9}{12}$$
$$E[Y^3] = 1^3\frac{1}{12}+2^3\frac{6}{12}+3^3\frac{5}{12} = \frac{184}{12}$$
$$E[X^2Y^3] = 1^22^2\frac{4}{12}+1^23^3\frac{5}{12} = \frac{167}{12}$$

#12 어느 집단 100명의 신장 X, 체중 Y의 분포는 다음과 같다. 이 집단에서 임의로 한 명을 뽑았을 대 그 사람의 신장을 X, 체중을 Y라 하자.

(1) $f_{X|Y}(X=160|Y=50)$

(2) 체중이 50일 때 신장의 기댓값은?

(3) 신장이 170이라면 체중의 기댓값은?

(4) 체질량계수의 기댓값은?

(5) $X=x$ 일 때 Y^2 의 조건부 기댓값은?

(6) 조건부 분포의 분산은?

Y \ X	150	160	170
50	10	5	0
60	5	30	10
70	0	20	20

(1) $f_{X|Y}(X=160|Y=50)$

$$= \frac{f_{X,Y}(X=160,\ Y=50)}{f_Y(Y=50)} = \frac{5}{15} = \frac{1}{3}$$

(2) 체중이 50일 때 신장의 기댓값은?

$$\frac{150\times10+160\times5}{15} = 150\frac{10}{15}+160\frac{5}{15} = 150f_{X|Y}(150|50)+150f_{X|Y}(160|50) = E(X|Y) = \sum_x xf(X|Y=y)$$

(3) 신장이 170이라면 체중의 기댓값은?

$$\frac{60\times10+70\times20}{30} = 60\frac{10}{30}+70\frac{20}{30} = 60f_{Y|X}(60|170)+70f_{Y|X}(70|170) = E(Y|X) = \sum_y yf(Y|X=x)$$

(4)체질량계수의 기댓값은?

체질량계수 = $\frac{1}{26}x^3\sqrt{y}$

$$E\left[\frac{1}{26}x^3\sqrt{y}\right] = \frac{1}{26}150^2\sqrt{50}\,\frac{10}{100}+\frac{1}{26}160^2\sqrt{50}\,\frac{5}{100}+\cdots+\frac{1}{26}170^2\sqrt{70}\,\frac{20}{100}$$

(5) $X=x$ 일 때 Y^2 의 조건부 기댓값은?

$$E(Y^2\,|\,X=x) = \sum_y y^2 f_{Y|X}(y\,|\,X=x)$$

(6) 조건부 분포의 분산은?

$$V(Y|X=x) = E(Y^2|X=x) - [E(Y|X=x)]^2$$

X와 Y가 독립이면 $E(X|Y=y)=E(x)$ 이 성립한다.

#13 조건부분포 연속일 때

$$f(x,y)=e^{-y}, \qquad 0<x<y<\infty$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_x^{\infty} e^{-y}dy = [-e^{-y}] = 0+e^{-x} = e^{-x}$$

$$f(Y|X) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{x-y}$$

$$E(Y|X=x) = \int_{-\infty}^{\infty} yf_{Y|X}(y|X=x)dy = \int_x^{\infty} ye^{x-y}dy = e^x\int_x^{\infty} ye^{-y}dy = x+1$$

$$V(Y|X=x) = E(Y^2|X=x) - [E(Y|X=x)]^2 = \int_x^{\infty} y^2e^{x-y}dy - (x+1)^2 = 1$$

#14 확률변수의 합

이항분포

$X\sim B(n_1,p)$, $Y\sim B(n_2,p)$ 이고 독립이면 $X+Y\sim B(n_1+n_2,p)$

포아송분포

$X\sim POI(\lambda_1)$, $Y\sim POI(\lambda_2)$ 이고 독립이면 $X+Y\sim POI(\lambda_1+\lambda_2)$

정규분포

$X_i\sim N(\mu_i,\sigma_i^2)$ 이고 독립이면 $\sum_{i=1}^k X_i \sim N(\sum_{i=1}^k \mu_i, \sum_{i=1}^k \sigma_i^2)$

표준정규분포의 제곱

$Z_i\sim N(0,1)$ 이고 독립이면 $\sum_{i=1}^k Z_i \sim \chi^2(k)$

#15 X, Y의 결합밀도함수가 $f_{X,Y}(x,y)=\begin{cases}\frac{1}{8}(x+y), & 0\le x,y\le 2\\ 0 & \text{그 외}\end{cases}$ 일 때 $P(|X-Y|>1)$ 의 값은 얼마인가?

$|X-Y|>1$

$X-Y>1$ or $X-Y<-1$ $\qquad P(|X-Y|>1)=\int_1^2\int_{Y+1}^2\frac{1}{8}(x+y)dxdy+\int_1^2\int_0^{Y-1}\frac{1}{8}(x+y)dxdy=\frac{1}{4}$

$X>Y+1$ or $X<Y-1$

#16 이변량확률변수 X,Y의 결합확률밀도함수가 다음과 같다.

$$f(x,y)=x+y,\qquad 0<x,y<1$$

(1) $f(y|X=0.5)$

(2) $E(Y|X=0.5)$

$f_X(x)=\int_0^1(x+y)dy=\left[xy+\frac{1}{2}y^2\right]=x+\frac{1}{2},\qquad f(X=0.5)=1$

$f(y|X=0.5)=\frac{f(0.5,y)}{f(X=0.5)}=\frac{0.5+y}{1}=y+\frac{1}{2}$

$E(Y|X=0.5)=\int_0^1 y(y+\frac{1}{2})dy=\int_0^1(y^2+\frac{1}{2}y)dy=\left[\frac{1}{3}y^3+\frac{1}{4}y^2\right]=\frac{7}{12}$

#17 이변량확률변수 X, Y의 결합밀도함수가 다음과 같다.

$$f(x,y)=3x+1,\qquad 0\le x,y,\qquad x+y<1$$

(1) $f_X(x)$, $f_Y(y)$

$f_X(x)=\int_0^{1-x}(3x+1)dy=[3xy+y]=3x(1-x)+(1-x)=(3x+1)(1-x)$

$f_Y(y)=\int_0^{1-y}(3x+1)dx=\left[\frac{3}{2}x^2+x\right]=\frac{3}{2}(1-y)^2+(1-y)$

#18 확률변수 X, Y의 결합확률밀도함수가 다음과 같다. 이 때 $f_X(x)$, $f_Y(y)$ 를 구하라. 또한 이 함수가 결합확률밀도함수임을 보이시오.

$$f(x,y)=e^{-(x+y)},\qquad x>0,\qquad y>0$$

$f_X(x)=\int_0^\infty e^{-(x+y)}dy=\left[-e^{-(x+y)}\right]=e^{-x}$

$f_Y(y)=\int_0^\infty e^{-(x+y)}dx=\left[-e^{-(x+y)}\right]=e^{-y}$

$\int_0^\infty\int_0^\infty e^{-(x+y)}dxdy=1$

#19 다음 이변수 함수가 확률변수 X와 Y의 결합밀도함수임을 보이시오.

$$f(x,y)=2,\qquad 0\le y\le x<1$$

$\int_0^1\int_y^1 2\,dxdy=\int_0^1[2x]dy=\int_0^1(2-2y)dy=\left[2y-y^2\right]=2-1=1$

$f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{2}{2-2y}=\frac{1}{1-y}$

$f_X(x)=\int_0^x 2dy=[2y]=2x$

$f_{Y|X}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{2}{2x}=\frac{1}{x}$

$f_Y(y)=\int_y^1 2dx=[2x]=2-2y$

#20 $f(x,y)$ 가 다음과 같을 때 k를 구하라.

$$f_{X,Y}(x,y) = \begin{cases} k(x+y), & x=0,1,2 \quad y=0,1,2 \\ 0 & \text{그 외} \end{cases}$$

$k(0+0)+k(0+1)+k(0+2)+k(1+0)+\ +k(1+1)+\ +k(1+2)+k(2+0)+k(2+1)+k(2+2) = 18k = 1$

$k = \frac{1}{18}$

#21 $f(x,y) = e^{-(x+y)}$, $x,y \geq 0$ 일 때 $P(X>2 | Y>1)$ 를 구하라.

$P(X>2,\ Y>1) = \int_1^{\infty}\int_2^{\infty} e^{-(x+y)}dxdy = e^{-3}$

$P(Y>1) = \int_1^{\infty} f_Y(y)dy = \int_1^{\infty}\int_0^{\infty} f_{X,Y}(x,y)dxdy = e^{-1}$

$P(X>2 | Y>1) = \frac{P(X>2,\ Y>1)}{P(Y>1)} = \frac{e^{-3}}{e^{-1}} = e^{-2}$

#22 $f(x,y) = \begin{cases} x+y, & 0<x,y<1 \\ 0, & \text{그 외} \end{cases}$ 일 때 공분산과 상관계수를 구하여라.

$f_X(x) = \int_0^1 (x+y)dy = x+\frac{1}{2}$

$E(X) = \int_0^1 x f_X(x)dx = \int_0^1 x^2+\frac{1}{2}x\ dx = \frac{7}{12}$

$f_Y(y) = \int_0^1 (x+y)dx = y+\frac{1}{2}$

$E(Y) = \int_0^1 y f_Y(y)dy = \int_0^1 y^2+\frac{1}{2}y\ dy = \frac{7}{12}$

$E(XY) = \int_0^1\int_0^1 xy f_{X,Y}(x,y)dxdy = \frac{1}{3}$

$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{49}{144} = \frac{-1}{144}$

$\rho_{XY} = \frac{Cov(X,Y)}{Var(X)Var(Y)}$